Markov Decision Processes in Finance and Dynamic Options.

Manfred Schäl
Inst. Angew. Math., Univ. Bonn, Wegelerstr.6, $D-53115$ Bonn

Draft from May 24, 2000

Dedicated to Professor Dr. Karl Hinderer

Abstract

In this paper a discrete-time Markovian model for a financial market is chosen. The fundamental theorem of asset pricing relates the existence of a martingale measure to the no-arbitrage condition. It is explained how to prove the theorem by stochastic dynamic programming via portfolio optimization. The approach singles out certain martingale measures with additional interesting properties. Furthermore, it is shown how to use dynamic programming to study the smallest initial wealth $x*$ that allows for super-hedging a contingent claim by some dynamic portfolio. There, a joint property of the set of policies in a Markov decision model and the set of martingale measures is exploited. The approach extends to dynamic options which are introduced here and are generalizations of American options.

1 Introduction and Summary

In discrete time $n = 0, 1, ..., N$ a financial market is studied which is free of arbitrage opportunities but incomplete. In the market $1+d$ assets can be traded. One of them with price process $\{B_n,\ 0 \leq n \leq N\}$ is called the bond or savings account and is assumed to be nonrisky. The other d assets are called stocks and are described by the d-dimensional price process $\{S_n,\ 0 \leq n \leq N\}$. An investor is considered whose attitude towards risk is specified in terms of a utility function U. A dynamic portfolio is specified by a policy ϕ. The investor's objective is to maximize the expected utility of the discounted terminal wealth $X_N^\phi(x)$ when starting with an initial wealth x .

Utility optimization is now a classical subject (Bertsekas [1], Hakansson [16]). The present paper is intended as a bridge between Markov decision processes (MDPs) and modern concepts of finance. It is assumed the reader knows some basic facts about MDPs, but knowledge about Mathematical Finance is not needed. In particular, we here are interested in using the no-arbitrage condition for the construction of a particular martingale measure by use of the optimal policy $\phi*$. A condition close to the no-arbitrage condition was already used by Hakansson [16]. Martingale measures are used for option pricing. The paper makes use of an approach how to base option pricing on the optimal solution $\phi*$ to the portfolio optimization problem. This approach is also explained by Davis [4].

In a one-period model, the construction of a martingale measure by use of an optimization problem is easily explained. Let us assume $d = 1$. If the action (portfolio) $a \in \mathbf{R}$ is chosen then the discounted terminal wealth of the portfolio has the form $x + a \cdot R$ where R can be interpreted as a return. If $a = a*$ is optimal then one has $\partial E[U(x + a \cdot R)]/\partial a = 0 = E[U'(x + a \cdot R) \cdot R]$ for $a = a*$. Upon defining a new measure Q by $dQ = const \cdot U'(x + a * \cdot R)dP$, one obtains $\int\ R\ dQ = E_Q[R] = 0$ which is the martingale property. Since the martingale property is a local one both in time and in space, martingale measures can be constructed by local optimization problems (see Rogers 1994) whereas we will apply global (dynamic) optimization. This has the advantage that the resulting martingale measures

have interesting interpretations. The case where the utility function U is only defined for positive values is treated in [38], [39].
A second interesting problem concerns values x for the initial wealth that allow for super-hedging discounted contingent claims $\breve{X}$ by some policy ϕ, *i.e.* $X_N^{\phi}(x) \geq \breve{X}$. The smallest value $x*$ coincides with the maximal expectation of $\breve{X}$ under (equivalent) martingale measures. The proof can make use of dynamic programming by exploiting an analogy between the set of all policies in a stochastic dynamic programming model and the set of martingale measures. A similar problem can be considered for American options where an optimal stopping time has to be chosen. It is well-known that an optimal stopping problem can be considered as a special stochastic dynamic programming problem. Therefore it is natural from the point of view of Markov decision theory to generalize the concept of an American option. This is done in the present paper and the generalization is called a dynamic option which has interesting applications.

2 The financial market

On the market an investor can observe the prices of $1 + d$ securities at the dates $n = 0, 1, ..., N$ where N is the *time horizon.* One of the securities is a *bond* (or savings account) with *interest rates* r_n, $1 \leq n \leq N$. It is essential for the theory that the interest rates for borrowing and lending are assumed to be the same. The *bond price process* is defined by
(2.1) $B_n := (1 + r_1) \cdot ... \cdot (1 + r_n)$, $0 \leq n \leq N$, where $B_0 = 1$.
Here we assume that $\{B_n\}$ is a deterministic process. If $\{B_n\}$ is given as the initial term structure, then the interest rates r_n can be computed by (2.1). We will have $r_n \geq 0$, but we only need that $1 + r_n > 0$. The other d securities are called stocks. The evolution of the *stock prices* will be modeled by a d-dimensional stochastic process $\{S_n, n = 0, 1, ..., N\}$ where S_0 is deterministic. There the components S_n^k of S_n, $1 \leq k \leq d$, are assumed to be positive. One may also think of a foreign exchange market where S_n^k is the exchange rate for a foreign currency. Besides the savings account, the investor has d accounts for different foreign exchanges. The value of these accounts as well as that of the savings account may be negative which can be interpreted as a loan.
The information about the market at time n including the observed stock prices will be represented by a Markov chain $(I_n, 0 \leq n \leq N)$ on some probability space (Ω, P) where I_n takes on values in some space E_n, $0 \leq n \leq N$, where $I_0 = i_0$ is a given constant and hence $E_0 = \{i_0\}$. In order to avoid integrability and measurability problems we assume here:
(2.2) E_n is finite, $1 \leq n \leq N$, hence $(w.l.o.g.)\Omega$ is finite and $P[\{\omega\}] > 0$, $\omega \in \Omega$.
For any vector-valued process $\{Z_n\}$, we define the backward increment by $\Delta Z_n := Z_n - Z_{n-1}$. Further, we write $\zeta^\top \cdot \xi$ for the inner product of $\xi, \zeta \in \mathbf{R}^d$.
The representation often becomes easier (see (2.6) below) if one considers the *discounted stock price process* $\breve{S}_n = (\breve{S}_n^1, ..., \breve{S}_n^d)$ defined by
(2.3) $\breve{S}_n^k := S_n^k / \; B_n$, $k = 1, ..., d$, $n = 0, ..., N$.
The *relative risk process* $\{R_n = (R_n^1, ..., R_n^d)$, $1 \leq n \leq N\}$ (Karatzas & Kou [22]) is defined by
(2.4) $1 + R_n^k := 1 + \Delta\breve{S}_n^k / \breve{S}_{n-1}^k = \frac{1}{1+r_n}\{1 + \Delta S_n^k / \; S_{n-1}^k\}$
where $\{\Delta S_n^k / S_{n-1}^k$, $1 \leq n \leq N\}$ is the *return process* corresponding to $\{S_n^k$, $0 \leq n \leq N\}$ (Pliska [31]§3.2). Then we get
(2.5) $\breve{S}_n^k = \breve{S}_{n-1}^k \cdot (1 + R_n^k) = S_0^k \cdot (1 + R_1^k) \cdot ... \cdot (1 + R_n^k)$.

Since the investor can observe S_n and R_n, it is natural to assume that S_n and R_n are known if the history $I_0, ..., I_n$ is known [*i.e.* that S_n and R_n are adapted in the sense explained in section 3]. In order to get a Markovian structure, we assume that R_n is a function $R_n = \rho_n(I_{n-1}, I_n)$ for some function ρ_n on $E_{n-1} \times E_n$. Important examples are $I_n = S_n$ or $I_n = R_n$. The latter example $I_n = R_n$ is sometimes convenient for a model where $R_1, ..., R_N$ are independent random variables as in many papers on portfolio optimization. An example for independent random variables $R_1, ..., R_N$ is the so-called *Binomial model*; there one has $d = 1$ and $r_n = r$ (independent of n). It is defined by two numbers δ and u such that $0 < \delta < 1 + r < u$ and $S_n = (1 + r) \cdot (1 + R_n) \cdot S_{n-1}$ is either equal to $u \cdot S_{n-1}$ or $\delta \cdot S_{n-1}$. If we choose $I_n = R_n$, then we get $E_n = \{\iota/(1 + r) - 1;\ \iota = \delta, u\}$ (which is independent of n). However, sometimes [*when* considering so-called contingent claims $X = f(S_N)$ in section 3] one should choose $I_n = S_n$ and then $E_n = \{u \cdot i,\ i \in E_{n-1}\} \cup \{\delta \cdot i,\ i \in E_{n-1}\}$. This example shows that it is useful to let E_n depend on n.

At each time n, the investor will obtain the market information $I_n = i$ and observe the discounted value x_n of his portfolio. Then he will decide about the amount ϕ_n^k invested in stock k during $(n, n + 1]$, *i.e.* , ϕ_n^k/S_n^k denotes the number of shares the investor holds during $(n, n+1]$. The decision may depend on the present information (i_n, x_n). A dynamic portfolio is here described by a (Markov) *policy* ϕ which is given by a sequence of functions $\phi = \{\phi_n,\ 0 \le n < N\}$ where ϕ_n is a mapping from $E_n \times \mathbf{R}$ to $\mathbf{R}^d$. In particular, one allows for negative amounts ϕ_n^k, *i.e.* one allows for short selling of stocks. In the case of a foreign exchange market one can think of a negative ϕ_n^k as a loan. Given the *initial wealth* x, the amount η_n invested in the bond in $[n, n + 1)$ is then specified by ϕ according to the following *budget equation* where we write $\underset{\sim}{1}$ for the d-dimensional vector with every tcol
component equal to 1.

(2.6) $\eta_0 + \phi_0^\top \cdot \underset{\sim}{1} = x,$ tcol

$\eta_n + \phi_n^\top \cdot \underset{\sim}{1} = \eta_{n-1} \cdot (1 + r_n) + \sum_{k=1}^d \phi_{n-1}^k \cdot S_n^k/S_{n-1}^k$ tcol

$= (1 + r_n) \cdot [\eta_{n-1} + \phi_{n-1}^\top \cdot (\underset{\sim}{1} + R_n)],\ 1 \le n < N.$ tcol

The investor can choose any new portfolio ϕ_n at time n for the stocks. This decision is then compensated by the savings account. In particular, this means that no funds are added to or withdrawn from the wealth of the portfolio at any time. All changes in wealth are due to capital gains (appreciation of stocks and interest from the bond). Then the policy is called *self-financing*.

Again it is convenient to work with the discounted wealth of the portfolio in place of the wealth itself. The *discounted wealth process* $\{X_n^\phi(x)\}$ is given through

(2.7) $X_n^\phi(x) := [\eta_n + \phi_n^\top \cdot \underset{\sim}{1}]/B_n = [\eta_{n-1} + \phi_{n-1}^\top \cdot (\underset{\sim}{1} + R_n)]\ /\ B_{n-1}\ ,\ 0 \le n < N,$ tcol

$X_N^\phi(x) := [\eta_{N-1} + \phi_{N-1}^\top \cdot (\underset{\sim}{1} + R_N)]\ /\ B_{N-1}\ .$ tcol

At time N there is no rebalancing of the portfolio. Obviously we have

(2.8) $X_n^\phi(x) = X_{n-1}^\phi(x) + \phi_{n-1}^\top \cdot R_n\ .$

It is also usual to describe a portfolio by the numbers $\xi_n^k := \phi_n^k/S_n^k$ of shares the investor holds. In the case where the wealth is positive it can be useful to describe a portfolio by the proportions $\pi_n^k := \phi_n^k/X_n^\phi(x)$ of the wealth which are invested in the stocks. But here it will be convenient to use the description of a portfolio by the amounts ϕ_n^k.

In the present approach we have $E_n \times \mathbf{R}$ as state space of the underlying Markov decision process and we tried to explain why it is useful to let E_n depend on n. It is well-known

that one can transform the present non-stationary Markovian model to a stationary model by including the time-parameter in the state (Feinberg [9]); but we will stick to the non-stationary setting. The action space is $\mathbf{R}^d$ and all actions $a \in \mathbf{R}^d$ are admissible at each time and state. Given the initial wealth x, the history $I_0, ..., I_n$, and the decisions $\phi_0, ..., \phi_{n-1}$, we know the discounted value $X_n^\phi(x) = x_n$ of the portfolio. Therefore, one could dispense with x_n as part of the state. However, the present choice (i_n, x_n) of the state makes the model Markovian. In the case (mostly considered in the literature) when $R_n = I_n$ and the $R_1, ..., R_N$ are independent, one can even dispense with i_n as part of the state and just choose x_n as the state at time n.
As underlying basic process we choose the process $\{I_n\}$ representing the informations about the market. There are transition probabilities $p_{ij}(n)$, $i \in E_{n-1}$, $j \in E_n$, where $p_{ij}(n)$ specifies the conditional probability $P[I_{n+1} = j | I_n = i]$ given the present value $I_n = i$ [*and* the past $(i_0, ..., i_{n-1})$]. It is a particular feature that the probability measure P describing the dynamics does not depend on the policy. This will be important when considering other artificial probability measures Q on Ω in the next section. As explained above, the state process for the Markov decision process is $\{(I_n, X_n^\phi(x))\}$. The distribution of the state process does depend on the policy as usual. Therefore, the measure P describing the dynamics of $\{I_n\}$ is here more appropriate than the policy depending (strategic) measure describing the dynamics of $\{(I_n, X_n^\phi(x))\}$ on the canonical space of trajectories of length N. If $\{I_n\}$ is a non-Markovian process, it can be transformed to the present non-stationary Markovian model by choosing the history space $E_0 \times ... \times E_n$ as new space $\tilde{E}_n$.

3 No-arbitrage and martingale measures

In order to use the important concept of martingales *w.r.t.* to the given information structure , we have to use conditional expectations $\zeta(i_1, ..., i_n) := E[Z | I_1 = i_1, ..., I_n = i_n]$ defined on $E_1 \times ... \times E_n$ and $E[Z|\sigma(I_1, ..., I_n)]$ defined on Ω for a random variable Z. As usual, we use the convention that $E[Z|\sigma(I_1, ..., I_n)](\omega) := \zeta(I_1(\omega), ..., I_n(\omega))$, $\omega \in \Omega$.
Then a real-valued stochastic process $\{Z_n, 0 \le n \le N\}$ is a *martingale* if Z_n is *adapted* to the information structure, *i.e.*, Z_n is a function of $(I_1, ..., I_n)$ where Z_0 is a constant and if

$E[\Delta Z_n | \sigma(I_1, ..., I_{n-1})] = 0$, $1 \le n \le N$.

It is necessary to consider further probability measures Q on Ω which are *equivalent* (to the given physical probability measure P); *i.e.*, $Q[\{\omega\}] > 0$, $\omega \in \Omega$. We write $E_Q[Z] = \int Z\, dQ$ for the expectation of the random variable Z under Q whereas $E[Z]$ is the expectation of the random variable Z under P as usual.

Definition 1. *Q is called a* **martingale measure** *iff $\{\check{S}_n\}$ forms a martingale under Q, i.e.,*

(3.1a) $E_Q[\Delta \check{S}_n^k | \sigma(I_1, ..., I_{n-1})] = 0$, $1 \le k \le d$, $1 \le n \le N$.

Obviously (3.1a) is equivalent to

(3.1b) $E_Q[R_n^k | \sigma(I_1, ..., I_{n-1})] = 0$, $1 \le k \le d$, $1 \le n \le N$.

Then we set

(3.2) $\mathcal{Q} := \{Q;\ Q$ is an equivalent martingale measure$\}$.

Equivalent martingale measures are used for the valuation of contingent claims (Harrison & Kreps [16]). A *contingent claim* is a random variable X that represents the time N payoff from a 'seller' to a 'buyer'. In most instances the random variable X can be taken to be some function of an underlying stock price, and so contingent claims are examples of what are called *derivative securities*. In the present framework, it is natural to assume that X is a function of I_N and we recall that we may choose $I_n = S_n$ as an example. A typical example is given by a *European call option*. At time $n = 0$ the buyer signs a contract which gives him the option to buy, at a specified time N (called the *maturity date* or *expiration date*), one share of stock 1 at a specified price χ, called the *exercise price*. At maturity, if the price S_N^1 of stock 1 is below the exercise price, the contract is worthless to the buyer; on the other hand, if $S_N^1 > \chi$, the buyer can exercise his option (*i.e.*, buy one share at the preassigned price χ) and then sell the share immediately in the market for S_N^1. This contract is thus equivalent to a payment of $X = (S_N^1 - \chi)^+$ (the contingent claim) at maturity. Given the price dynamics of the securities, one tries to determine the prices of such a contingent claim X. A classical answer in the sense of Huygens and Bernoulli to the question of a *fair price* π for the option would rely on the concept of a fair game and would then be $\pi = E[X/B_N]$. There is a different answer in the present situation where the seller can hedge X by investing in the stocks For example if S_N^1 is high, then the buyer's and the seller's gain is high.

A contingent claim X is said to be *attainable* if there exists a policy ϕ and some $x \in \mathbf{R}$ such that $\check{X} = X_N^\phi(x)$ for $\check{X} := X/B_N$. The corresponding policy ϕ is said to replicate X, since the (nondiscounted) wealth $B_N \cdot X_N^\phi(x)$ of the dynamic portfolio ϕ at time N replicates X. This is the case if $\chi = 0$, *i.e.* $X = S_N^1$, in the example above; then one can choose $x = S_0^1$ and $\phi_n^1 \equiv S_n^1$, $\phi_n^k \equiv 0$, $k \neq 1$, for all n. It is clear that for attainable contingent claims with $\check{X} = X_N^\phi(x)$ a fair price is x. The buyer and the seller could instead have invested the wealth x in such a way as to replicate the payoff of the contingent claim. Now from the martingale property one immediately gets

(3.3) $E_Q[X_N^\phi(x)] = x$ for $Q \in Q$.

Since a price system should be a linear functional on the space of all contingent claims X, *i.e.* of all random variables X on Ω (Harrison & Kreps [16]), it becomes clear that

(3.4) $\pi_Q(X) := E_Q[\check{X}]$ for $Q \in Q$

is a candidate for a fair price. The market is said to be *complete* if each contingent claim is attainable. An example is the Binomial model for the case $d = 1$. For complete markets option pricing is no problem. However, there are only very few examples of complete discrete-time markets (Harrison & Pliska [17], Jacod & Shiryaev [20]). In continuous time, the Black-Scholes model is an important example of a complete market. The question of the existence of some $Q \in Q$ is strongly connected with the following *no-arbitrage condition*:

(NA) for $1 \leq n \leq N$ and any portfolio ϕ_{n-1} depending on $(I_1, ..., I_{n-1})$:

$\phi_{n-1}^\top \cdot R_n \geq 0$ implies $\phi_{n-1}^\top \cdot R_n = 0$.

There (NA) means that if there is a chance that $\phi_{n-1}^\top \cdot R_n > 0$, then there is also a chance that $\phi_{n-1}^\top \cdot R_n < 0$. A portfolio ϕ_{n-1} with "$\phi_{n-1}^\top \cdot R_n \geq 0$" and "$\phi_{n-1}^\top \cdot R_n > 0$ for some ω" is called an *arbitrage opportunity*. This definition of arbitrage is standard in the literature where we here need not care about null sets. Such an arbitrage opportunity represents a riskless source of generating profit, strictly greater than the profit from the bond. In order to present (NA) in full generality, we used a non-Markovian dynamic portfolio ϕ. Now we can present one of the most important results for financial markets:

Fundamental Theorem of Asset Pricing. *There exists an equivalent martingale mea-*

sure, i.e., Q is not empty, if and only if the no-arbitrage condition (NA) *holds.*

Proofs for general spaces Ω can be found in Dalang et al [3], Schachermayer [35], Rogers [32], Jacod & Shiryaev [20]. It is easy to see that (NA) is necessary for the existence of some $Q \in Q$; therefore we can give the proof here: Let be $Q \in Q$. From $\phi_{n-1}^{\top} \cdot R_n \geq 0$ we conclude that $\phi_{n-1}^{\top} \cdot R_n = 0 \;\; Q-a.s.$ since $E_Q[\phi_{n-1}^{\top} \cdot R_n] = 0$. As Q is equivalent we have $\phi_{n-1}^{\top} \cdot R_n = 0$ everywhere. In this paper it will be proved in §6 by use of dynamic programming that (NA) is also sufficient for the existence of some $Q \in Q$. Since the martingale property is a local one both in time and in space, martingale measures can also be constructed by local optimization problems. This was done by Rogers [32] whereas we will apply global (dynamic) optimization. From now on we use the following assumption:
Assumption 1. (NA) *holds.*
Since in discrete time the assumption of completeness is a severe restriction, we are forced to consider general incomplete markets. From the so-called second fundamental theorem of asset pricing it is known (Harrison & Pliska [17], Jacod & Shiryaev [20]) that in incomplete markets with the (NA)-condition one has several choices of equivalent martingale measures (from the convex set Q). Thus in incomplete markets, no preference independent pricing of contingent claims is possible. The approach of this paper, which makes use of dynamic programming, has another interesting consequence about identifying a certain price. There option pricing is based on an optimal solution to the portfolio optimization problem. This approach is also explained by Davis [4]. Suppose a utility function U on $\mathbf{R} \times E_N$ is given where $U(x,i)$ is strictly concave and differentiable in x with partial derivative $U'(x,i)$. The investor's objective is to maximize the expected utility of terminal wealth and the maximum utility is given by
(3.5) $V_0(x) = \sup_{\phi} E[U(X_N^{\phi}(x), I_N)] = E[U(X_N^{\phi*}(x), I_N)]$
where $\phi*$ is a solution to the optimization problem. In §6 it will be explained that it may be useful to let the utility depend on I_N. Consider a contingent claim X made available for trading with purchase price π. One can ask the question whether the maximum utility in (3.5) can be increased by the purchase (or short-selling) of an option described by X. In order to find a fair price $\hat{\pi}$ for X one can use the following argument: $\hat{\pi}$ is a fair price for the contingent claim if diverting a little of the investor's initital wealth into the option at time zero has a neutral effect on the investor's achievable utility. For $N = 1$ this concept was also explained by Merton [27]§7.7). More precisely, consider the discounted terminal wealth at time N of an investor who follows the dynamic portfolio $\phi*$ which is optimal for his initial wealth x and who then diverts the amount $\xi \in \mathbf{R}$ to purchase ξ/π shares of the option with price π and discounted contingent claim $\breve{X} = X/B_N$:

$$X_N^{\phi*}(x-\xi) + \frac{\xi}{\pi} \cdot \breve{X} = x - \xi + X_N^{\phi*}(0) + \frac{\xi}{\pi} \cdot \breve{X} = X_N^{\phi*}(x) + \xi \cdot \{\frac{1}{\pi}\breve{X} - 1\} \; .$$ tcol

Then the expected utility is
(3.6) $g(\phi*, \xi, \pi, x) = E[U(X_N^{\phi*}(x) + \xi \cdot \{\frac{1}{\pi}\breve{X} - 1\}, I_N)]$ tcol
where $\sup_{\phi} g(\phi, 0, \pi, x) = V_0(x) = g(\phi*, 0, \pi, x)$. Then one obtains
(3.7) $\frac{\partial}{\partial \xi} g(\phi*, \xi, \pi, x)\Big|_{\xi=0} = E[U'(X_N^{\phi*}(x), I_N) \cdot \{\frac{1}{\pi}\breve{X} - 1\}]$. tcol
Further, $g(\phi*, \xi, \pi, x)$ is concave in ξ and strictly concave except for the less interesting case where $\breve{X}$ agrees with π and hence X is attainable.

Now choose $\pi = \hat{\pi}$ such that $\frac{\partial}{\partial \xi} g(\phi*, \xi, \hat{\pi}, x)\big|_{\xi=0} = 0$ and $g(\phi*, \xi, \hat{\pi}, x)$ is hence maximal for $\xi = 0$. Thus $\hat{\pi}$ is not too high otherwise the investor would like to sell short the option; and $\hat{\pi}$ is not too low otherwise the investor would like to purchase the option. This leads to the formula tcol

(3.8) $\hat{\pi} = E[U'(X_N^{\phi*}(x), I_N) \cdot \check{X}] \;/\; E[U'(X_T^{\phi*}(x), I_N)] = \int \check{X} dQ_x^U$.

The relation (3.8) leads to the interesting result that in our situation the option pricing formula of Davis is given by the martingale measure Q_x^U which is constructed in §5. A similar result was found by Karatzas & Kou [22] Theorem 7.4, who studied a diffusion model extensively. The case where U is defined only for positive values is treated in [38], [39].

Again consider an option with discounted contingent claim $\check{X} := X/B_N$. In the second part of the paper we are interested in

(3.9) $x* := \inf\{x \in \mathbf{R};\; X_N^{\phi}(x) \geq \check{X}$ for some $\phi\}$.

Thus we look for values x of the initial wealth that allow for *super-hedging* X by some dynamic portfolio ϕ. Then a duality result holds: The smallest value $x*$ will be shown to coincide with the maximal expectation of $\check{X}$ under equivalent martingale measures.

Theorem 2. $x* = \sup_{Q \in Q} E_Q[\check{X}]$.

The quantity $x*$ is called upper price, hedging price, upper bound for the fair price, selling price, or arbitrage upper bound. It is known (see *e.g.* [37] 1.16) that $E_Q[\check{X}] < x*$ for each $Q \in Q$ unless X is attainable. In view of Theorem 2, a price $\pi = E_Q[\check{X}]$ (for some $Q \in Q$) thus offers no arbitrage opportunity to the seller in the sense that $X_N^{\phi}(\pi) \geq \check{X}$, $X_N^{\phi}(\pi) \neq \check{X}$ for some ϕ.

A quantity similar to $x*$ will also be considered for *American options* (Karatzas [21]) where an optimal stopping time can be chosen by the buyer. Moreover, a generalization of American options is introduced here which is called *dynamic option*. This is a natural generalization from the point of stochastic dynamic programming where it is known how to imbed an optimal stopping problem in a Markov decision problem.

4 The one-period model

In this section we will study the case $N = 1$. We write $R := R_1$ and

(4.1) $\sum := \{R(\omega);\, \omega \in \Omega\}$

for the *support* of R. Furthermore, $\mathcal{L}$ is the smallest linear space in $\mathbf{R}^d$ containing $\sum$, *i.e.* $\mathcal{L}$ is the smallest linear space L in $\mathbf{R}^d$ such that $P[R \in L] = 1$. Then it is easy to show that (NA) is equivalent *to* :

$(NA)_1$ for all $a \in \mathcal{L} \setminus \{0\} : P[a^\top \cdot R < 0] > 0$.

We start from a utility function $U(x, i)$ depending on two variables and we use the following properties:

Assumption 3. $U \in \mathbf{F}_N$ *where*

$\mathbf{F}_n := \{f : \mathbf{R} \times E_n \mapsto \mathbf{R}$ such that $x \mapsto f(x, i)$ is strictly concave and differentiable with derivative $f'(x, i)$; furthermore, $f'(-\infty, i) > 0$ and $f'(+\infty, i) \leq 0\}$.

¿From the concavity we know that $U'(x, i)$ is decreasing in x. In a one-period model, a policy ϕ specifies just one portfolio $a \in \mathbf{R}^d$. We obtain from (2.6) and (2.7) that then

(4.2) $X_1^{\phi}(x) := x + a^\top \cdot R$.

We want to maximize the expected utility:
(4.3) $v(x, a) := E[U(x + a^\top \cdot R,\, I_1)]$ for $a \in \mathbf{R}^d$;

$V(x) := \sup_{a \in \mathbf{R}^d} v(x, a)$ for $x \in \mathbf{R}^d$.

Lemma 4. (a) *If* Γ *denotes the orthogonal projection on* $\mathcal{L}$ *, then* $a^\top \cdot R = (\Gamma a)^\top \cdot R$;
(b) $(x, a) \mapsto v(x, a)$ *is continuous*;
(c) *for each* $x \in \mathbf{R}$, $a \mapsto v(x, a)$ *attains the maximum on* $\mathbf{R}^d$ *where the maximum point can be chosen in* $\mathcal{L}$ *.*

Proof. The parts (a) and (b) are obvious. For part (c) one can use the methods of Leland [26], Bertsekas [1] or Rogers [32]. By (a) we can restrict attention to $a \in \mathcal{L}$. We have

$$\frac{1}{\lambda} \cdot [v(x, \lambda \cdot a) - v(x, 0)]$$ tcol

$$= E[\, D(\lambda, a^\top \cdot R, x, I_1) \cdot 1_{\{a^\top \cdot R > 0\}}] + E[D(\lambda, a^\top \cdot R, x, I_1) \cdot 1_{\{a^\top \cdot R < 0\}}].$$

where $D(\lambda, y, x, i) := \frac{1}{\lambda} \cdot \{U(x + \lambda\, y, i) - U(x, i)\}$ is decreasing in λ both for $y > 0$ and for $y < 0$ because of the concavity of $U(\cdot, i)$. Further tcol
$D(\infty, y, x, i) := \lim_{\lambda \blacksquare \infty} D(\lambda, y, x, i) \le 0$ for $y > 0$ and $D(\infty, y, x, i) < 0$ for $y > 0$ by our assumption on U. Now, we get by use of the monotone convergence theorem:

$$\lim_{\lambda \blacksquare \infty} \frac{1}{\lambda} \cdot [v(x, \lambda \cdot a) - v(x, 0)]$$ tcol

$$= E[\, D(\infty, a^\top \cdot R, x, I_1) \cdot 1_{\{a^\top \cdot R > 0\}}] + E[D(\infty, a^\top \cdot R, x, I_1) \cdot 1_{\{a^\top \cdot R < 0\}}].$$

By (NA) we have for $a \in \mathcal{L} \setminus \{0\} : P[a^\top \cdot R < 0] > 0$ and thus:

$$\lim_{\lambda \blacksquare \infty} \frac{1}{\lambda} \cdot [v(x, \lambda \cdot a) - v(x, 0)] < 0 \text{ for } a \in \mathcal{L} \setminus \{0\}.$$ tcol

In particular, we have $\lim_{\lambda \blacksquare \infty} v(x, \lambda \cdot a) = -\infty$ for $a \in \mathcal{L} \setminus \{0\}$. Now the result follows (see Rockafellar [33] Theorems 27.1, 27.3, Bertsekas [1] Proposition 1, Rogers [32] Proposition 2.2). []

Lemma 5. (a) $v(x, a)$ *is strictly concave in* x *for fixed* a;
(b) $V(x)$ *is strictly concave in* x *and the maximum point of* $a \mapsto v(x, a)$ *in* $\mathcal{L}$ *is unique.*

Proof. Part (a) is obvious. For the proof of (b) choose $x, \tilde{x} \in \mathbf{R}$ and $\lambda, \tilde{\lambda} > 0$ such that $\lambda + \tilde{\lambda} = 1$. Further choose $a \in \mathcal{L}$ such that $V(x) = v(x, a)$ and similarly $\tilde{a}$ for $\tilde{x}$. We consider the cases (i) $\tilde{x} > x$ and (ii) $\tilde{x} = x$ and $a \neq \tilde{a}$. Then $P[x + a^\top \cdot R \neq \tilde{x} + \tilde{a}^\top \cdot R]] > 0$; otherwise $(a - \tilde{a})^\top \cdot R = \tilde{x} - x$ which contradicts (NA) in case (i) and contradicts $a - \tilde{a} \in \mathcal{L} \setminus \{0\}$ in case (ii). Now we obtain from the strict concavity of $U(\cdot, i)$: $\lambda \cdot V(x) + \tilde{\lambda} \cdot V(\tilde{x}) = \lambda \cdot v(x, a) + \tilde{\lambda} \cdot v(\tilde{x}, \tilde{a}) < v(\lambda \cdot x + \tilde{\lambda} \cdot \tilde{x}\ ,\ \lambda \cdot a + \tilde{\lambda} \cdot \tilde{a}) \le V(\lambda \cdot x + \tilde{\lambda} \cdot \tilde{x})$. In case (ii) we have a contradiction. []

Lemma 6. *For* $a \in \mathbf{R}^d$, $v(x, a)$ *is differentiable in* x *with derivative*

$v'(x, a) = E[U'(x + a^\top \cdot R,\, I_1)]$.

The proof is obvious since Ω is finite.

Theorem 7. $V(x)$ *is differentiable in* x *and* $V'(x) = v'(x, a*)$ *where* a* $= a * (x)$ *is the maximum point in* $\mathcal{L}$ *of the function* $\mathbf{R}^d \blacksquare a \mapsto v(x, a)$ *.*

Proof. From Lemmata 4 and 5 we know that a unique maximum point a* exists in $\mathcal{L}$. Now $V(x \pm \blacksquare) - V(x) \ge v(x \pm \blacksquare, a*) - v(x, a*)$. Hence $v'_+(x, a*) \le V'_+(x) \le V'_-(x) \le v'_-(x, a*)$ where $V'_\pm(x)$ and $v'_\pm(x, a*)$ denote the right and left derivatives, respectively. Now Lemma 6 applies. []

Lemma 8. *For $x > 0$, $v(x, a)$ is partially differentiable in $a \in \mathbf{R}^d$ with partial derivatives*
$\partial_k v(x, a) = E[U'(x + a^\top \cdot R, I_1) \cdot R^k]$.
Again, the proof is obvious since Ω is finite.

Theorem 9. *Let be $x \in \mathbf{R}$, then $E[U'(x + a*^\top \cdot R, I_1) \cdot R^k] = 0$ for $1 \le k \le d$ where a* is the unique maximum point in $\mathcal{L}$ of the function $\mathbf{R}^d \ni a \mapsto v(x, a)$.*
Proof. The function $a \mapsto v(x, a)$ is partially differentiable by Lemma 8. Therefore we know for the maximum point a* that $\partial_k v(x, a*) = 0$ for $1 \le k \le d$. []

Corollary 10. *If $U'(x, i) > 0$ for all x and i or more generally if $U'(x + a*^\top \cdot R, I_1) > 0$ on Ω, then*
the probability measure Q defined by $dQ = const \cdot U'(x + a*^\top \cdot R, I_1)dP$ is an equivalent martingale measure.

5 The multi-period model
We remind the reader that $R_n = \rho_n(I_{n-1}, I_n)$ for some function ρ_n. Now we need the support of R_n given $I_{n-1} = i$ defined by $\sum_n(i) := \{\rho_n(i, j); j \in E_n, p_{ij}(n) > 0\}$, $i \in E_{n-1}$, $n \ge 1$. $\mathcal{L}_n(i)$ is the smallest linear space in $\mathbf{R}^d$ containing $\sum_n(i)$, *i.e.* $\mathcal{L}_n(i)$ is the smallest linear space L in $\mathbf{R}^d$ such that $P[\ R_n \in L | I_{n-1} = i] = 1$. Then for $\sum$ and $\mathcal{L}$ as defined in §4 we have $\sum = \sum_1(i_0)$ and $\mathcal{L} = \mathcal{L}_1(i_0)$.
We may assume (*w.l.o.g.*) that the no-arbitrage condition also holds locally (see Dalang et al. [3] Lemma 2.3 , Pliska [31](3.22), Schäl [37]§2) , *i.e.*,
$(NA)*$ "$a^\top \cdot \sigma \ge 0 \forall \sigma \in \sum_n(i)$" implies "$a^\top \cdot \sigma = 0 \forall \sigma \in \sum_n(i)$" , $i \in E_{n-1}$, $a \in \mathbf{R}^d$.
The condition (NA)* just means:
"$P[\ a^\top \cdot R_n \ge 0 | I_{n-1} = i] = 1$" implies "$P[\ a^\top \cdot R_n = 0 | I_{n-1} = i] = 1$" , $i \in E_{n-1}$, $a \in \mathbf{R}^d$.
As in §4, we will use the assumption $U \in \mathbf{F}_N$ for the utility function U. We recall that by (2.8): $X_N^\phi(x) = X_n^\phi(x) + \sum_{m=n+1}^N \phi_{m-1}^\top \cdot R_m$ and we now define:
(5.1) $v_n(x, i, \phi) = E[\ U(x + \sum_{m=n+1}^N \phi_{m-1}^\top \cdot R_m\ ,\ I_N) | I_n = i\ ]$,
$V_n(x, i) = \sup_\phi v_n(x, i, \phi)$,
$v_N(x, i, \phi) = V_N(x, i) = U(x, i)$ for all ϕ.
There $v_n(x, i, \phi)[V_n(x, i)]$ is the [*maximal*] expected utility of the terminal wealth given the market information i and the discounted wealth x at time n. Since $I_0 = i_0$ is fixed, we can set:
(5.2) $v_0(x, \phi) = v_0(x, i_0, \phi)$, $V_0(x) = V_0(x, i_0)$.
Obviously we have an N-stage Markov decision model with no running costs and terminal reward U. Let us now introduce the well-known *reward operator*:
(5.3) $T_n^a f(x, i) := E[f(x + a^\top \cdot R_{n+1}\ ,\ I_{n+1}) | I_n = i]$ for any $f : \mathbf{R} \times E_{n+1} \mapsto \mathbf{R}$, $a \in \mathbf{R}^d$, $i \in E_n$.
We can express $T_n^a f(x, i)$ in (5.3) by the transition matrix $(p_{ij}(n))$ according to
(5.4) $T_n^a f(x, i) = \sum_{j \in E_{n+1}} p_{ij}(n) \cdot f(x + a^\top \cdot \rho_{n+1}(i, j), j)$
Thus for fixed i, $T_n^a f(x, i)$ can be expressed by an ordinary expectation and we can use the results of §4. The following equation is sometimes called *fundamental equation* (Dynkin & Yushkevich [6]) and follows from Fubini's theorem (Hinderer [18] Lemma 11.1).
(5.5) $v_n(x, i, \phi) = T_n^{\phi_n(x,i)} v_{n+1}(x, i, \phi)$, $n \ge 0$.

Now we can give the well-known *optimality equation* (see Hinderer 1970, Theorems 14.4, 18.4):
(5.6) $V_n(x,i) = \sup_{a\in\mathbf{R}^d} T_n^a V_{n+1}(x,i) = \sup_{a\in\mathcal{L}_{n+1}(i)} T_n^a V_{n+1}(x,i)$, $n \geq 0$.

Proposition 11. *Let be $f \in \mathbf{F}_{n+1}$ and $0 \leq n < N$. Then:*
(a) $(x,a) \mapsto T_n^a f(x,i)$ *is continuous;*
(*b*) *there exists a unique measurable function $\delta : \mathbf{R} \times E_n \mapsto \mathbf{R}^d$ such that*
$\delta(x,i) \in \mathcal{L}_{n+1}(i)$ *and*
$T_n^{\delta(x,i)} f(x,i) = \max_{a\in\mathbf{R}} T_n^a f(x,i) =: F(x,i)$;
(*c*) *$F(x,i)$ is strictly concave in x.*
Proof. Part (a) and part (*b*) follow from Lemma 4. The remaining properties follow from Lemma 5. []
Finally, we will use another assumption which will be discussed in §6 for special cases.
Assumption 12. *For $V_n'(x,i) = \partial V_n(x,i)/\partial x$ we have: $V_n'(-\infty,i) > 0$ and $V_n'(+\infty,i) \leq 0$, $0 \leq n < N$.*
The existence of $V_n'(x,i)$ follows from:
Lemma 13. (a) *$V_n(x,i)$ is differentiable in x with derivative $V_n'(x,i)$.*
(*b*) *$V_n(x,i) \in \mathbf{F}_n$ for $0 \leq n \leq N$.*
Proof. Part (a) can be proved as Theorem 7. By assumption we know that $V_T = U$ is in $\mathbf{F}_N$. Upon using the optimality equation (5.6), this statement follows by backward induction from Assumption 12 and Proposition 11c. []
Now we obtain from (5.5), (5.6), Proposition 11, and Lemma 13 by the usual arguments of dynamic programing (Feinberg [9]) :
Theorem 14. *There exists a unique policy $\phi* = (\phi_n^*)$ such that for $0 \leq n < N$:* tcol
$\phi_n^*(x,i) \in \mathcal{L}_{n+1}(i)$ *and* tcol
$T_n^{\phi_n^*(x,i)} V_{n+1}(x,i) = \max_{a\in\mathbf{R}^d} T_n^a V_{n+1}(x,i) = V_n(x,i)$,
$v_n(x,i,\phi*) = V_n(x,i)$.
Proof. We just have to define $\phi_n = \delta$ as in Proposition 11b where $f = V_{n+1}$. []

Proposition 15. *For some $x > 0$, let $\phi*$ be the optimal policy of Theorem 14. Then*
$V_0'(x) = E[U'(X_N^{\phi*}(x), I_N)]$.
Proof. By induction we are going to prove
(5.7) $V_0'(x) = E[V_n'(X_n^{\phi*}(x), I_n)]$, $1 \leq n \leq N$.
For $n = 1$ we have
$V_0(x) = \max_{a\in\mathcal{L}_1(i_0)} E[V_1(x + a^\top \cdot R_1, I_1)]$ where $\phi_0(x,i_0)$ is a maximum point.
In view of Lemma 13, $V_1(x,i)$ satisfies the Assumption 3 for $U(x,i)$ (for $N = 1$). With the help of Lemma 6 and Theorem 7, we can conclude:
$V_0'(x) = v_0'(x, \phi_0^*(x,i_0)) = E[V_1'(x + \phi_0^*(x,i_0)^\top \cdot R_1,\, I_1)]$ tcol
$= E[V_1'(X_1^{\phi*}(x), I_1)]$.
Hence (5.7) holds for $n = 1$. Now assume that (5.7) holds for $n-1$. For fixed $i \in E_{n-1}$ we have
$V_{n-1}(x,i) = \max_{a\in\mathcal{L}_n(i)} E[V_n(x + a^\top \cdot R_n, I_n) | I_{n-1} = i]$
where $\phi_{n-1}^*(x,i)$ is a maximum point. In view of Lemma 13, $V_n(x,i)$ satisfies the Assumption 3 for $U(x,i)$ (for $N = n$). As above we obtain: tcol

$V_{n-1}'(x,i) = E[V_n'(x+\phi_{n-1}^*(x,i)^\top \cdot R_n,\ I_n)|I_{n-1}=i].$ tcol

Then by (2.8) : $X_n^{\phi*}(x) = X_{n-1}^{\phi*}(x) + \phi_{n-1}^{*\top} \cdot R_n$ tcol

Since $X_{n-1}^{\phi*}(x)$ is a function of $(I_1,...I_{n-1})$, it follows that

(5.8) $V_{n-1}'(X_{n-1}^{\phi*}(x), I_{n-1})$

$= E[V_n'(X_{n-1}^{\phi*} + \phi_{n-1}^*(x,i)^\top \cdot R_n,\ I_n)|\sigma(I_1,...,I_{n-1})]$ tcol

$= E[V_n'(X_n^{\phi*}, I_n)|\sigma(I_1,...,I_{n-1})]$.

By assumption we get

$V_0'(x) = E[V_{n-1}'(X_{n-1}^{\phi*}(x), I_{n-1})] = E[E[V_n'(X_n^{\phi*}, I_n)|\sigma(I_1,...,I_{n-1})]] = E[V_n'(X_n^{\phi*}(x)), I_n)]$

and (5.7) follows. Now we can use (5.7) for $n = N$ and finally obtain the result. []

Theorem 16. *Assume that U' is positive or more generally that $U'(X_N^{\phi*}(x), I_N)$ is positive where $\phi*$ is the optimal policy of Theorem 14. Define for some x the process $\{Z_n,\ 0 \le n \le N\}$ by*

$Z_n := V_n'(X_n^{\phi*}(x), I_n)$, *in particular* $Z_0 = V_0'(x)$, $Z_N = U'(X_N^{\phi*}(x), I_N)$, *then:*

(a) *One obtains an equivalent martingale measure Q_x^U on Ω by*

$$Q_x^U[\{\omega\}] = \frac{1}{V_0'(x)} Z_N(\omega)P[\{\omega\}] = (Z_N(\omega)/Z_0)P[\{\omega\}].$$ tcol

(*b*) *$\{Z_n/Z_0,\ 0 \le n \le N\}$ is a martingale under P and is called the density process of dQ_x^U/dP .*

(*c*) *If we write $E_x^U[...]$ for the expectation under Q_x^U , then for any function f_n*

$E_x^U[f_n(I_1,...,I_n)|\sigma(I_1,...,I_{n-1})] = E[f_n(I_1,...,I_n)\cdot Z_n|\sigma(I_1,...,I_{n-1})]\ /\ Z_{n-1}.$

Proof. If $U'(X_N^{\phi*}(x), I_N)$ is positive then we know from Proposition 15 that Q_x^U is indeed a probability measure which is equivalent to P. Now part (*b*) immediately follows from (5.8) and part (*c*) is a consequence of Bayes' rule. From Theorem 9 we conclude that

(5.9) $E[V_n'(x+\phi_{n-1}^*(x,i)^\top \cdot R_n,\ I_n)\cdot R_n^k|I_{n-1}=i] = 0$ for $1 \le k \le d$ tcol

since $\phi_{n-1}^*(x,i)$ is the unique maximum point in $\mathcal{L}_n(i)$ of the function tcol

$\mathbf{R}^d \ni a \mapsto E[V_n'(x+a^\top \cdot R_n,\ I_n)|I_{n-1}=i]$. ¿From (5.9) we obviously obtain as in proof of (5.8):

$E[R_n^k \cdot Z_n|\sigma(I_1,...,I_{n-1})] = 0$, $1 \le k \le d$.

Finally from (*c*) we get

(5.10) $E_x^U[R_n^k|\sigma(I_1,...,I_{n-1})] = 0$, $1 \le n \le N$, $1 \le k \le d$.[]

6 Applications

In this section we will show that all assumptions are satisfied for the case where

(6.1) $U(x,i) = -e^{-\gamma x}\cdot L(i)$ for some $\gamma > 0$ and some function $L : E_N \mapsto (0,\infty)$.

A classical example is $U(x,i) = -\frac{1}{\gamma}e^{-\gamma x}$ (or $U(x,i) = \frac{1}{\gamma}(1-e^{-\gamma x})$). If one wants to consider the utility of the terminal wealth itself rather that than the discounted terminal wealth, then one can choose $U_B(x,i) := U(B_N \cdot x, i)$ which has the same form (6.1). A further interesting example was studied by Grandits & Rheinländer [15]. They look for a policy ϕ^0 such that tcol

(6.2) $E[exp\{\gamma\cdot(\check{X} - X_N^{\phi^0}(x))\}] = \inf_\phi$.

Instead of super-hedging the contingent claim X one tries to choose ϕ^0 for the given initial wealth x such that the expected loss is minimized. There the positive values of $\check{X} - X_N^{\phi^0}(x)$

are given more weight than the negative values. Of course, this is the view of the seller. If one chooses $I_N = S_N$ and $\check{X} = f(S_N)$ as in the case of a European call option, then this problem is included in the framework of (6.1) by choosing $L(i) := \frac{1}{\gamma}\exp\{\gamma \cdot f(i)\}$. There tcol
are some recent papers treating the problem $E[\ell(\{\check{X} - X_N^{\phi^0}(x)\}^+)] = \inf_\phi$ (see Runggaldier & Zaccaria [34] for further references and an approach via dynamic programming in the case of transaction costs).

For convenience, we only will consider the case $\gamma = 1$.

Lemma 17. *In the situation* (6.1) *with* $\gamma = 1$, *one has*:

(a) *Assumption* 3 *is satisfied and* $U'(x,i) = e^{-x} \cdot L(i)$ *is positive*;

(*b*) $V_n(x,i) = -e^{-x} \cdot L_n(i)$ *for some function* $L_n : E_n \mapsto (0,\infty)$;

(*c*) *for the optimal policy* $\phi*$ *of Theorem* 14, $\phi_n^*(x,i)$ *does not depend on* x; tcol

(*d*) *Assumption* 12 *is satisfied.*

Proof. Part (a) is obvious and part (*d*) follows from part (*b*). Part (*b*) is obvious for $n = N$. Now assume that (*b*) holds for $n+1$. Then according to Theorem 14
$V_n(x,i) = \max_{a\in\mathcal{L}_{n+1}(i)} -\sum_{j\in E_{n+1}} p_{ij}(n)\exp\{-x - a^\top \cdot \rho_{n+1}(i,j)\}\cdot L_{n+1}(j) =: -e^{-x}\cdot L_n(i)$
where the maximum point $\phi_n(i)$ is indeed independent of x and

(6.3) $L_n(i) = \min_{a\in\mathcal{L}_{n+1}(i)} \sum_{j\in E_{n+1}} p_{ij}(n)\exp\{-a^\top \cdot \rho_{n+1}(i,j)\}\cdot L_{n+1}(j)$. []

Consider the case where $R_1, ..., R_N$ are independent and $I_n = R_n$. Then both $p_{ij}(n)$ and $\rho_{n+1}(i,j)$ are independent of i. From the proof of Lemma 17 we conclude that in that case $\phi_n^*(x,i) =: \phi_n^*$ neither depends on x nor on i. tcol

Definition 18. *In the situation* (6.1) *with* $\gamma = 1$, *let* $\phi*$ *be the optimal policy of Theorem 14 which is independent of the initial wealth* x *according to Lemma* 17. *Then* $X_n^* := X_n^{\phi*}(0)$ tcol
, $n = 0, ..., N$ *is called the* **optimal wealth process**.

Lemma 19. *In the situation* (6.1) *with* $\gamma = 1$, *the density process of Theorem* 16 *is given by*

$Z_n/Z_0 = \exp\{-X_n^*\} \cdot L_n(I_n)/L_0(i_0)$, $0 \leq n \leq N$, *where* $Z_N/Z_0 = \exp\{-X_N^*\} \cdot L(I_N)/L_0(i_0)$. tcol

Proof. We have $Z_n = V_n'(x + X_n^*, I_n) = e^{-x}\cdot\exp\{-X_n^*\}\cdot L_n(I_n)$, thus $Z_0 = e^{-x}\cdot L_0(i_0)$ tcol
and $Z_N = e^{-x}\cdot\exp\{-X_N^*\}\cdot L(I_N)$. [] tcol

Now we obtain from Theorem 16 the following result:

Theorem 20. *Let* $\{X_n^*\}$ *be the optimal wealth process in the situation* (6.1) *with* $\gamma = 1$ tcol
and let L_n *be defined by* (6.3) *where* $L_N = L$. *Then* Q^U *defined by*

$Q^U[\{\omega\}] := \exp\{-X_N^*(\omega)\}\cdot L(I_N(\omega))/L_0(i_0)P[\{\omega\}]$ *is a martingale measure.* tcol

Upon choosing $L(i) = 1$, for example, we obtain the existence of a special martingale measure and the Fundamental Theorem of §3 is proved where (NA) was used in the proof of Lemma 4c. Then the resulting martingale measure Q^E has the form:

(6.4) $Q^E[\{\omega\}] = const \cdot \exp\{-X_N^*(\omega)\}P[\{\omega\}]$. tcol

There Q^E is the unique solution of the minimum problem where the relative entropy $I(Q,P)$ of Q *w.r.t.* P has to be minimized (Fritelli [13], Grandits [14]). Q^E may also be considered as a multi-stage *Esscher transform* of P whereas in Bühlmann et al. [2] the Esscher transform is used locally for each time and each history in order to construct a martingale measure.

Another important example is the following:
(6.5) $U(x,i) = L^{(0)}(i) + 2\ L^{(1)}(i) \cdot x - L^{(2)}(i) \cdot x^2$
for some function $L^{(k)} : E_N \mapsto \mathbf{R}$, $k = 0,1,2$ where $L^{(2)}$ is positive.
An interesting special case is the best hedging policy for hedging the contingent claim X. One looks for a policy ϕ^o such that
(6.6) $E[\{\check{X} - X_N^{\phi^o}(x)\}^2] = \inf_\phi$.
Thus, instead of super-hedging X one tries to choose ϕ^o for the given initial wealth x such that the expected quadratic loss is minimized. Now, in contrast to (6.2), the positive values of $\check{X} - X_N^{\phi^o}(x)$ have the same weight as the negative values. This is fair both from the point of view of the seller and the buyer. If one again chooses $I_N = S_N$ and $\check{X} = f(S_N)$, then this problem is included in the framework of (6.5) by choosing $L^{(0)}(i) = f(i)^2$, $L^{(1)}(i) = f(i)$, and $L^{(2)}(i) := 1$.
Now, one can look for that initial wealth $\hat{x}$ such
(6.7) $\inf_\phi E[\{\check{X} - X_N^{\phi}(\hat{x})\}^2] = \min_x \inf_\phi E[\{\check{X} - X_N^{\phi}(x)\}^2]$.
The optimal value $\hat{x}$ is called *fair hedging price* in [36] and *approximation price* by Schweizer [40] for the contingent claim X. Whereas [36] relies on stochastic dynamic programming, Schweizer [40] uses martingale methods.

Lemma 21. *In the situation* (6.5) *one has*:
(a) *Assumption* 3 *is satisfied* ;
(b) $V_n(x,i) = L_n^{(0)}(i) + 2\ L_n^{(1)}(i) \cdot x - L_n^{(2)}(i) \cdot x^2$
for some function $L_n^{(k)} : E_n \mapsto \mathbf{R}$, $k = 0,1,2$, *where* $L_n^{(2)}$ *is positive*;
(c) *Assumption* 12 *is satisfied.*

Proof. Part (a) is obvious and part (c) follows from part (b). Part (b) is obvious for $n = N$. Now assume that (b) holds for $n+1$. Then according to Theorem 14

$$V_n(x,i) = \max_{a\in\mathbf{R}^d} \sum_{j\in E_{n+1}} p_{ij}(n+1)[L_{n+1}^{(0)}(j) + 2\ L_{n+1}^{(1)}(j) \cdot \{x + a^\top \cdot \rho_{n+1}(i,j)\}$$
$$-L_{n+1}^{(2)}(j)\{x + a^\top \cdot \rho_{n+1}(i,j)\}^2]$$
$$= \tilde{L}^{(0)}(i) + 2\tilde{L}^{(1)}(i)x - \tilde{L}^{(2)}(i)x^2 + L*(x,i) - \hat{L}(x,i)$$

where $\tilde{L}^{(k)}(i) := \sum_j p_{ij}(n+1)L_{n+1}^{(k)}(j)$, $k = 0,1,2,$

$$L*(x,i) := -\min_{a\in\mathbf{R}^d} \sum_j p_{ij}(n+1) \cdot L_{n+1}^{(2)}(j)\{L_{n+1}^{(1)}(j)/L_{n+1}^{(2)}(j) - x - a^\top \cdot \rho_{n+1}(i,j)\}^2,$$
$$\hat{L}(x,i) := \sum_j p_{ij}(n+1)L_{n+1}^{(2)}(j)\{L_{n+1}^{(1)}(j)/L_{n+1}^{(2)}(j) - x\}^2.$$

Now $\hat{L}$ has the desired form. This is also true for $L*$ as follows from [36] Proposition 7.3 for the case $d = 1$ and is easily extended to the case $d \geq 1$. []

The problem with this example (6.5) is that $U'(x,i) = 2\ L^{(1)}(i) - 2\ L^{(2)}(i) \cdot x$ is not positive everywhere and also $U'(X_N^{\phi*}(x), I_N)$ need not to be positive everywhere. Therefore the measure Q_x^U is in general only a signed measure and is called signed martingale measure because it has all other properties of a martingale measure. In Schweizer [40], [41] the relations to the *mean-variance frontier*, the *variance-optimal* and the *minimal martingale measures* are explained. Schweizer [41] also considers the general d-dimensional case in a general semi-martingale framework which includes the discrete-time case as special case. Motoczynski [28] studies the discrete-time d-dimensional setting. Grandits [14] constructs the *p-optimal martingale measure* as a generalization of the variance-optimal martingale measure by replacing L^2 by L^p for some $p > 1$. For the construction Grandits also uses methods of dynamic programming similar to those explained in § 5.

Example. The multi-dimensional Binomial model.

For convenience we assume $T = 1$ and write $R^k := R^k_1$, $1 \le k \le d$. We consider the model where

R^k takes on values in $\{-\alpha^k, \beta^k\}$ where $\alpha^k, \beta^k > 0$.

We assume *w.l.o.g.* that $I_1 = R_1 = R$. For $d = 1$ the model is complete and is also called Cox-Ross- Rubinstein model. The unique martingale measure $P*$ is given by $P*[R^1 = i] = q^*_1(i)$ where tcol

$q^*_k(-\alpha^k) := \beta^k/(\alpha^k + \beta^k)$, $q^*_k(\beta^k) := \alpha^k/(\alpha^k + \beta^k)$. tcol

For $d > 1$, the model is no longer complete since R takes on 2^d values where $2^d > 1 + d$ (Jacod & Shiryaev [20]). It is easy to see that $Q*[R = (i^1, ..., i^d)] = \Pi^d_{k=1} q^*_k(i^k)$ defines a tcol martingale measure for $d \ge 1$ which seems to be a natural one if the components R^k of R are independent under P. Infact in that case this measure coincides with Q^E as defined in (6.4) but is different from the so-called *minimal martingale measure* which is obtained by the (discrete-time) *Girsanov transformation* and which coincides with the *variance-optimal martingale measure* for $T = 1$ (Schweizer [40]). $Q*$ also differs from the martingale measure obtained by the so-called *numéraire portfolio* (Korn & Schäl [23]). We want to prove:

$Q* = Q^E$ if the components R^k of R are independent under P.

We first observe that $P* = Q* = Q^E$ if $d = 1$ since then there is only one martingale measure. Hence for $\inf_{a\in\mathbf{R}} E[exp\{a \cdot R^k\}] = E[exp\{a^*_k \cdot R^k\}] =: A_k$ we know that $q^*_k(i) =$ tcol $\exp\{a^*_k \cdot i\}/\ A_k$. Now $\inf_{a\in\mathbf{R}^d} E[exp\{x + a^\top R\}] = e^x \cdot \inf_{a\in\mathbf{R}^d} \Pi^d_{k=1} E[exp\{a^k \cdot R^k\}]$ tcol $= e^x \cdot \Pi^d_{k=1} \inf_{a\in\mathbf{R}} E[exp\{a \cdot R^k\}] = e^x \cdot \Pi^d_{k=1} A_k$ by independence. Therefore we obtain

$Q^E[\{(i^1, ..., i^d)\}] \ /\ P[\{(i^1, ..., i^d)\}] = \exp\{x + \sum^d_{k=1} a^*_k \cdot i^k\}/(e^x \cdot \Pi^d_{k=1} A_k)$ tcol

$= \Pi^d_{k=1} \exp\{a^*_k \cdot i^k\}/\ A_k = \Pi^d_{k=1} q^*_k(i^k)$. tcol

The result immediately extends to $T > 1$ if one assumes that $R^1_1, ..., R^d_1, ..., R^1_T, ..., R^d_T$ are independent. Motoczynski & Stettner [29] consider super-hedging (see next sections) in the multi- dimensional multi-period Cox-Ross-Rubinstein model. []

7 Super-hedging, the one-period model.

In this section we restrict attention to the case $N = 1$ and study values x for the initial wealth that allow for super-hedging a given discounted contingent claims $\breve{X}$ by some policy ϕ such that $X^\phi_N(x) \ge \breve{X}$. As in §4 we set $R := R_1$. Then we have to solve the problem:

(7.1) $x* = \inf\{x \in \mathbf{R};\ \exists a \in \mathbf{R}^d s.th.\ x + a^\top \cdot R(\omega) \ge \breve{X}(\omega)\ ,\ \omega \in \Omega\}$.

Since $\omega \in \Omega$ is finite, one obtains a linear program by introducing:

(7.2) $V := \{(x, a) \in \mathbf{R}^{1+d};\ 1 \cdot x + \sum_k a^k \cdot R^k(\omega) \ge \breve{X}(\omega)\ ,\ \omega \in \Omega\}$.

Then (7.1) can be written as the linear program:

(P) $\min\{x;\ (x, a) \in V\}$.

Then the dual program is:

(D) $\max\{\sum_{\omega\in\Omega} q(\omega) \cdot \breve{X}(\omega)\ ;\ q \in Q_1\}$ where

(7.3) $\hat{Q}_1 := \{q \in \mathbf{R}^{|\Omega|};\ q(\omega) \ge 0,\ \omega \in \Omega,\ \sum_\omega q(\omega) = 1,\ \sum_\omega q(\omega) \cdot R^k(\omega) = 0,\ 1 \le k \le d\}$.

$\hat{Q}_1$ can be identified with set of martingale measures in the case $N = 1$. A useful characterization is given in Pliska [31] *p.* 59. The set of equivalent martingale measures is

(7.4) $Q_1 := \{q \in \hat{Q} \; ; \, q(\omega) > 0, \, \omega \in \Omega\}$.
Obviously, V is not empty; one only has to choose x large enough. From the Fundamental Theorem in §3 we know that Q_1 and hence $\hat{Q}_1$ is not empty under the no-arbitrage condition.
Now the duality theorem of linear programming applies and we obtain (Pliska [31] *p.* 27):

Proposition 21. $x* = \min\{x \in \mathbf{R}; \, \exists a \in \mathbf{R}^d s.th. \; x + a^\top \cdot R(\omega) \geq \breve{X}(\omega) \; , \, \omega \in \Omega\}$
$= \max_{q \in \hat{Q}1} \sum_{\omega \in \Omega} q(\omega) \cdot \breve{X}(\omega) = \sup_{q \in Q_1} \sum_{\omega \in \Omega} q(\omega) \cdot \breve{X}(\omega) =: \sup_{q \in Q_1} E_q[\breve{X}]$. tcol

The latter equality holds because Q_1 is dense in $\hat{Q}_1$. This is theorem 2 in the case $N = 1$. An immediate consequence is:

Corollary 22. $E_q[\breve{X}] \leq 0 \forall q \in Q_1$ *if and only if there exists some* $a \in \mathbf{R}^d$ *such that* $a^\top \cdot R \geq \breve{X}$.

8 Super-hedging, the multi-period model.

In this section we consider the general case $N \geq 1$ and we may assume (NA)* as in §5. It will be convenient to work with a canonical probability space:
(8.1) $\Omega := \Omega_N$ where $\Omega_n := E_0 \times ... \times E_n$, ,
$I_n(\omega) = i_n$, $\omega_n := (i_0, ..., i_n)$ for $\omega = (i_0, ..., i_N) \in \Omega$.
We need the following characterization of Q([33], Korn & Schäl [23]).
Proposition 23. $Q \in Q$ *if and only if*
(8.2) $Q[\{\omega\}] = q_0(i_0; i_1) \cdot q_1(i_0, i_1; i_2) \cdot ... \cdot q_{N-1}(\omega_{N-1}; i_N)$
where $q_n(\omega_n; \cdot) \in Q_n(i_n)$ *for* $\omega = (i_0, ..., i_N)$ *and*
(8.3) $Q_n(i) := \{q; \, q : E_{n+1} \mapsto (0, 1); \, \sum_{j \in E_{n+1}} q(j) = 1 \; , \; \sum_{j \in E_{n+1}} q(j) \cdot \rho(i, j) = 0\}$, $i \in E_n$.
There is now an interesting relation to dynamic programming. One can look upon $Q_n(i)$ as a set of actions available at time n given the information i about the market at time n. Moreover, a function q_n can be considered as a function which selects for each given history ω_n an action $q_n(\omega_n; \cdot) \in Q_n(i_n)$. Then a policy is given by $(q_0, q_1, ..., q_{N-1})$ and defines a martingale measure through (8.2). In models with a more general space Ω, the situation is more complicated. But even then the set Q enjoys a property ([37] Lemma 1.8) which is known for the set of policies in dynamic programming and is called "to admit needle-like variation" by Fakeev [8], "stability" by Hinderer [18], and "product property" by Hordijk]19]. This relation to dynamic programming was used by El Karoui & Quenez [7] for diffusion models and by Naik & Uppal [30].
As mentioned above, it is natural from the point of view of Markov decision theory to generalize the concept of an American option. We call the generalization a *dynamic option* which has interesting applications. The buyer has to pay a premium x to the seller and then he can choose any policy δ according to a certain dynamic program. Then the policy δ implies some (non- discounted) cost or claim $X^\delta := B_N \cdot \breve{X}^\delta : \Omega \mapsto \mathbf{R}$ which the seller has to pay to the buyer at N. On the other side, the seller can invest according to a policy ϕ in the market, *i.e.* the seller is the investor. Now, we look for some initial wealth x such that
(8.4) for each policy δ of the buyer there is some policy $\phi(\delta)$ of the seller such that $X_N^{\phi(\delta)}(x) \geq \breve{X}^\delta$ on Ω,

i.e. such that $\breve{X}^\delta$ can be super-hedged by x and $\phi(\delta)$. In particular, we are again interested in the smallest value $x*$ such that (8.4) holds for $x = x*$.
We obtain the usual European option if there exists only one policy for the buyer. One obtains an American call option if the buyer may choose a stopping time $\tau : \Omega \mapsto \{0, ..., N\}$ and may exercise his option at time τ . Such a contract is equivalent to a payment $X^\tau = (S^1_\tau - \chi)^+$. Then one is interested in finding some initial wealth x and some policy ϕ such that for the discounted claim:
(8.5) $X^\phi_\tau(x) \geq \breve{X}^\tau = (\breve{S}^1_\tau - \chi/B_\tau)^+ = (S^1_\tau - \chi)^+/B_\tau$ for all stopping times τ.
We can write $X^\phi_\tau(x) = X^{\phi(\tau)}_N(x)$ where $\phi_n(\tau) := \phi_n$ for $n \leq \tau$ and $\phi_n(\tau) := 0$ for $n > \tau$. There, first the buyer decides to stop and then the seller decides to stop investing. Thus, the American option fits into the framework of a dynamic option and the investment policy of the seller indeed depends on τ which is not clear from (8.5) at first glance.
A necessary condition for (8.4) can be deduced from (3.3) (*i.e.* $E_Q[X^\phi_N(x)] = x$ for any $Q \in Q$). We obtain:
(8.6) $\bar{x} := \sup_{Q,\delta} E_Q[\breve{X}^\delta] \leq x* := \inf\{x;\ (8.4) \text{ holds }\}$.
Now our goal is to show that
(8.7) $\forall\delta\exists\phi(\delta)$ such that $X^{\phi(\delta)}_N(\bar{x}) \geq \breve{X}^\delta$.
Then $\bar{x} = x*$ and the infimum in the definition of $x*$ is attained.
First we describe a dynamic program for the buyer by the usual set up. Here it is useful to choose a model with no running costs and a terminal reward (claim) $\breve{X}$ which depends on the whole history at time N. [*In* such a framework one can model restrictions on the admissible actions by the choice of the terminal reward.] In order to describe general stopping times of the buyer by policies of the buyer we have to allow for non-Markovian policies.

(1) $\underline{\underline{A}}'_n$ is the *action space* of the buyer; tcol
(2) $H_n := E_0 \times A'_0 \times ... \times E_n$ is the *space of histories* of the buyer up to epoch n;
(3) a *policy* $\delta = (\delta_0, ..., \delta_{N-1})$ of the buyer is given by functions $\delta_n : \Omega_n \mapsto \underline{\underline{A}}'_n$; tcol
(4) $\breve{X} : H_N \mapsto \mathbf{R}$ is a *discounted claim function* where we write $\breve{X}^\delta(\omega) := \breve{X}(h^\delta_N(\omega))$ and $h^\delta_n(\omega_n) := (i_0, \delta_0(i_0), i_1, ..., \delta_{n-1}(\omega_{n-1}), i_n) \in H_n$.

At each time n the buyer can choose an action a'_n according to some policy δ and the seller will immediately be informed about the choice of a'_n at time n. The claim of the dynamic option is executed at time N; at the end the seller has to pay the amount $B_N \cdot \breve{X}^\delta(\omega)$ to the buyer.
For an American call option, one will have $\underline{\underline{A}}'_n := \{0, 1\}$ where '1' means 'to stop' and tcol '0' stands for 'to do nothing'. Moreover, one can choose $\breve{X}(i_0, a'_0, i_1, ..., a'_{N-1}, i_N) = (i_\tau - \chi)^+/B_\tau$ where
$\tau := \inf\{n;\ a_n = 1\}$ with $\inf \emptyset = N$. One can imagine that the claim $(i_\tau - \chi)^+$ at time τ is invested in the savings account up to time N; at N the amount $(i_\tau - \chi)^+ \cdot B_N/B_\tau$ is paid to the buyer including interest for the time between τ and N. Discounting then leads to $(i_\tau - \chi)^+/B_\tau$. Thus the American option can also be described by a claim executed at time N.
We will present another example.

Example from the bond market. Assume that at time $n = 0$ the buyer gives some capital y_0 to the seller. In that situation, an action a' consists in taking back the amount a'. Then

$\underline{\underline{A}}'_n = [0, \bar{a}_n]$ for some $\bar{a}_n \geq 0$ and for $n \geq 0$. tcol

Actually admissible actions a'_n at time n are only those with $a'_n \leq \alpha_n$ where

$\alpha_n := \bar{a}_n \wedge (y_0 - a'_1 - ... - a'_{n-1})$, $n < N$.

This can be modeled by either defining $\breve{X}(i_0, a'_0, i_1, ..., a'_{N-1}, i_N) = -\infty$ if some a'_n is not admissible or by setting $\breve{X}(i_0, a'_0, i_1, ..., a'_{N-1}, i_N) = \breve{X}(i_0, a'_0 \wedge \alpha_0, i_1, ..., a'_{N-1} \wedge \alpha_{N-1}, i_N)$. In the later case the actions a'_n and $a'_n \wedge \alpha_n$ are identified. []

A further modern example is provided by a *passport option* (see Delbaen & Yor [5], Shreve & Vecer [42] which allows the buyer to take (long or short) positions in a stock. If at N the buyer makes a benefit, he can keep it. Otherwise he does not have to pay for the losses.

The idea of our approach consists in writing $\bar{x} := \sup_{Q,\delta} E_Q[\breve{X}^\delta]$ as value of a dynamic program in a super-model with a super-player combining the buyer choosing δ (controlling the claim) and the market choosing the martingale measure Q (controlling the law of motion).

As before, the initial state i_0 is fixed. The data are

($\hat{1}$) $\hat{\underline{\underline{A}}}_n := \underline{\underline{A}}'_n \times \{q : E_{n+1} \mapsto (0,1), \sum_{i \in E_{n+1}} q(i) = 1\}$, $n \geq 0$, tcol

($\hat{2}$) $\hat{\underline{\underline{A}}}_n(i_n) = \hat{\underline{\underline{A}}}'_n \times Q_n(i_n)$ for $n \geq 0$, tcol

($\hat{3}$) $\hat{\delta} := (\delta, Q)$ where $\delta = (\delta_0, ..., \delta_{N-1})$, $Q = (q_0, ..., q_{N-1})$,

($\hat{4}$) $\hat{X} := \breve{X}$, $\hat{X}^\delta := \breve{X}^\delta$;

($\hat{5}$) the law of motion is given through Q as in Proposition 23.

A policy δ is called admissible at $h_n = (i_0, a'_0, ..., i_n)$ if $\delta_m(i_0, ...i_m) = a'_m$ for $m < n$. Then the value functions of the super-model are given by

(8.8) $\hat{V}_n(h_n) := \sup\{E_Q[\breve{X}^\delta | I_0 = i_0, ..., I_n = i_n];\ \delta$ is admissible at h_n, $Q \in Q\}$,

for $h_n = (i_0, a'_0, ..., i_n) \in H_n$.

Then $\hat{V}_0(i_0) := \bar{x}$.

An important tool will be the optimality equation (Hinderer [18], Theorem 14.4):

(8.9) $\hat{V}_n(h_n) = \sup_{a' \in \underline{\underline{A}}'_n, q \in Q_n(i_n)} \sum_{i \in E_{n+1}} q(i) \hat{V}_{n+1}(h_n, a', i)$,

which implies:

Lemma 24. $\sum_{i \in E_{n+1}} q(i)[\hat{V}_{n+1}(h_n, a', i) - \hat{V}_n(h_n)] \leq 0 \forall a' \in \underline{\underline{A}}'_n$, $q \in Q_n(\omega_n)$. tcol

Now we can use the results for the one-period model by use of the *standard reduction method* of Hinderer [18] *p.* 24. The idea consists in treating time n as new origin for fixed (h_n, a'). Then $Q_n(i_n)$ is the set of equivalent martingale measures for the new one-period model. Now Corollary 22 applies and we obtain:

Proposition 25. *For fixed* $(h_n, a') = (i_0, a'_0, ..., i_n, a')$ *there exists some* $\bar{\phi}_n(h_n, a')$ *such that*

$\hat{V}_{n+1}(h_n, a', i) - \hat{V}_n(h_n) \leq \bar{\phi}_n(h_n, a')^\top \cdot \rho_{n+1}(i_n, i) \forall i \in E_{n+1}$.

One can interpret $\bar{\phi}_n(h_n, a')$ as the decision of the seller at time n about the investment in the market after the buyer chose action a'. Therefore the super-model can be looked

upon as a stochastic game with complete information in the sense of Küenle [25]. Now we can construct the desired investment policy, which will be non-Markovian, as follows:
(8.10) $\bar{\phi}_n^\delta(\omega_n) := \bar{\phi}_n(h_n^\delta(\omega_n), \delta_n(\omega_n))$, $n \geq 0$, where $h_0^\delta(\omega_0) = \omega_0 = i_0$.
There $h_n^\delta(\omega_n)$ is defined above. Then one has by Proposition 24:
(8.11) $\hat{V}_{n+1}(h_n^\delta(\omega_n), \delta_n(\omega_n), i_{n+1}) - \hat{V}_n(h_n^\delta(\omega_n)) \leq \bar{\phi}_n^\delta(\omega_n)^\top \cdot \rho_{n+1}(i_n, i_{n+1})$.
Now we can prove our goal (8.7):

Theorem 26. $X_N^{\bar{\phi}(\delta)}(\bar{x}) \geq \check{X}^\delta$ *where* $\bar{\phi}(\delta) := (\bar{\phi}_n^\delta)$.

Proof. Choose some $h = h_N = (i_1, a_1', ..., i_N)$ with histories $h_n = (i_1, a_1', ..., i_n)$, $\omega = (i_1, ..., i_N)$, $\omega_n = (i_1, ..., i_n)$. Then by Proposition 25

$\hat{V}_N(h) - \hat{V}_0(i_0) = \sum_{n=0}^{N-1} [\hat{V}_{n+1}(h_{n+1}) - \hat{V}_n(h_n)]$
$\quad \leq \sum_{n=0}^{N-1} \bar{\phi}_n(h_n, a_n')^\top \cdot \rho_{n+1}(i_n, i_{n+1})$ where
$\quad \hat{V}_N(h) = \check{X}(h)$, $\hat{V}_0(i_0) = \bar{x}$ by (8.6) and (8.8).

Thus we have
$\quad \check{X}(h) \leq \bar{x} + \sum_{n=0}^{N-1} \bar{\phi}_n(h_n, a_n')^\top \cdot \rho_{n+1}(i_n, i_{n+1})$ for $h \in H_N$
which implies
(8.12) $X^\delta(\omega) \leq \bar{x} + \sum_{n=0}^{N-1} \bar{\phi}_n^\delta(\omega_n)^\top \cdot \rho_{n+1}(i_n, i_{n+1}) = X_N^{\phi(\delta)}(\bar{x})$. []

Lemma 24 can be looked upon as a supermartingale property under each martingale measure which was derived by means of dynamic programming in the same sense as in El Karoui & Quenez [7]. One can write (8.12) as
(8.13) $X^\delta(\omega) = X_N^{\phi(\delta)}(\bar{x}) + \sum_{n=1}^{N} c_n^\delta$ where
$\quad c_{n+1}^\delta := [\hat{V}_{n+1}(h_{n+1}^\delta(\omega_{n+1})) - \hat{V}_n(h_n^\delta(\omega_n))] - \bar{\phi}_n^\delta(\omega_n)^\top \cdot \rho_{n+1}(i_n, i_{n+1}) \geq 0$.
Then c_n^δ can be interpreted as a possible consumption at time n when hedging $\check{X}$.
Theorem 26 can be imbedded in the framework of a general optional decomposition theorem for supermartingales under each martingale measure. (Föllmer & Kabanov [10], Föllmer & Kramkov [11], Kramkov [24], Schäl [37]).
The analogy between the set of all policies in Markov decision model and the set of martingale measures can also be used to show that Markovian martingale measures are sufficient in the same sense as Markovian policies are sufficient in a Markovian environment ([37]).
Acknowledgement. I would like to thank the referee for very useful remarks. Morover, it is a pleasure to acknowledge the influence of my teacher Karl Hinderer who laid much of the foundation of the general non-Markovian theory of dynamic programming. Modern finance is another important area where the non-Markovian setting is natural. In particular, contingent claims of so-called exotic options provide interesting examples of general history-dependent reward and cost functions, respectively.

Bibliography

[1] D. Bertsekas, "Necessary and sufficient conditions for existence of an optimal portfolio", *J. Econ. Theory* **8**, 235-247, 1974.
[2] H. Bühlmann, F. Delbaen, P. Embrechts and A. Shiryaev, "On Esscher transforms in discrete finance models", *Astin Bulletin* **28**, 171 — 186, 1998.
[3] R.C. Dalang, A. Morton and W. Willinger, "Equivalent martingale measures and no-arbitrage in stochastic securities market models", *Stochastics and Stochastic Reports* **29**, 185 — 201, 1990.

[4] M.H.A. Davis, “Option pricing in incomplete markets”, in: *Mathematics of Derivative Securities*, ed: M.A.H. Dempster, S.R. Pliska, Cambridge Univ. Press, 216 — 226, 1997.
[5] F. Delbaen, and M. Yor, “Passport options”, Working paper, ETH Zürich, 1999.
[6] E.B. Dynkin and A.A. Yushkevich, *Controlled Markov Processes*, Springer, Berlin, 1997.
[7] N. El Karoui and M.C. Quenez, “Dynamic programming and pricing of contingent claims in an incomplete market”, *SIAM J. Control Optim.* **33**, 29-66, 1995.
[8] A.G. Fakeev, “Optimal stopping rules for processes with continuous parameter”, *Theory Probab. Appl.* **15**, 324-331, 1970.
[9] E.A. Feinberg, “Total reward criteria”, This volume, 2000.
[10] H. Föllmer and Yu.M. Kabanov, “Optional decomposition and Lagrange multipliers”, *Finance Stochast.* **2**, 69 — 81, 1998.
[11] H. Föllmer and D. Kramkov, “Optional decompositions under constraints”, *Probab. Theory Relat. Fields* **109**, 1 — 25, 1997.
[12] H. Föllmer and M. Schweizer, “Hedging of contingent claims under incomplete information”, in: *Applied Stochastic Analysis*, eds. M.H.A. Davis and R.J. Elliot, Gordon and Breach, London, 1990.
[13] M. Fritelli, “The minimal entropy martingale measure and the valuation problem in incomplete markets”, *Mathematical Finance*. To appear .
[14] P. Grandits, “The p-optimal martingale measure and its asymptotic relation with the minimal-entropy martingale measure”, *Bernoulli* **5**, 225-248, 1999.
[15] P. Grandits and Th. Rheinlaender “On the minimal entropy martingale measure”, Preprint, Techn. Univ. Berlin, Fb. Mathem. 1999.
[16] N. H. Hakansson, “Optimal entrepreneurial decisions in a completely stochastic environment”, *Management Science* **17**, 427-449, 1971.
[16] J.M. Harrison and D.M. Kreps, “Martingales and arbitrage in multiperiod securities markets”, *J. Economic Theory* **20**, 381-408, 1979.
[17] J.M. Harrison and S.R. Pliska, “Martingales and stochastic integrals in the theory of continuous trading”, *Stoch. Processes & Appl.* **11**, 215-260, 1981.
[18] K. Hinderer, *Foundations of non-stationary dynamic programming with discrete time-parameter*, Lecture Notes in Operations Research and Mathematical Systems **33** Berlin-Heidelberg-New York: Springer, 1970.
[19] A. Hordijk, *Dynamic Programming and Markov Potential Theory*, Amsterdam: Mathematical Centre Tracts **51**, 1974.
[20] J. Jacod and A. N. Shiryaev, “Local martingales and the fundamental asset pricing theorems in the discrete-time case”, *Finance Stochast.* **3**, 259-273, 1998.
[21] I. Karatzas, “On the pricing of American Options”, *Appl.Math. Optim.* **17**, 37-60, 1988.
[22] I. Karatzas and S.G. Kou, “On the pricing of contingent claims under constraints”, *Ann. Appl. Probab.* **6**, 321-369, 1996.
[23] R. Korn and M. Schäl, “On value preserving and growth optimal portfolios”, *Math. Methods Op. Res.* **50**, 189 — 218, 1999.
[24] D.O. Kramkov, “Optional decomposition of supermartingales and hedging of contingent claims in incomplete security markets”, *Probab. Theory Relat. Fields* **105**, 459 — 479, 1996.

[25] H.-U. Küenle, *Stochastische Spiele und Entscheidungsmodelle*, Teubner-Texte zur Mathematik, 89. Teubner, Leibzig, 1986.

[26] H. Leland, "On the existence of optimal policies under uncertainty", *J. Econ. Theory* **4** , 35-44, 1972.

[27] R.C. Merton, *Continuous-time Finance*, Blackwell, Cambridge, MA, 1990.

[28] M. Motoczynski, "Multidimensional variance-optimal hedging in discrete time model — a general approach", *Mathematical Finance* **10**, 243-258, 2000.

[29] M. Motoczynski and L. Stettner, "On option pricing in the multidimensional Cox-Ross-Rubinstein model", *Applicationes Mathematicae* **25**, 55-72, 1998.

[30] V. Naik and R. Uppal, "Minimum cost hedging in incomplete Markets", Univ. of British Columbia, Working paper, 1992.

[31] S.R. Pliska, *Introduction to Mathematical Finance*, Blackwell publisher, Malden, USA, Oxford, UK, 1997.

[32] L.C.G. Rogers, "Equivalent martingale measures and no-arbitrage", *Stochastics and Stochastic Reports* **51**, 41-49, 1994.

[33] R.T. Rockafeller, *Convex Analysis*, Princeton Univ. Press, Princeton, *N.J*, 1970.

[34] W. J. Runggaldier and A. Zaccaria, "A stochastic control approach to risk management under restricted information", *Mathematical Finance* **10**, 277-288, 2000.

[35] W. Schachermayer, "A Hilbert space proof of the fundamental theorem of asset pricing in finite discrete time", *Insurance: Mathematics and Economics* **11**, 249 — 257, 1992.

[36] M. Schäl, "On quadratic cost criteria for option hedging", *Math. Oper. Res.* **19**, 121-131, 1994.

[37] M. Schäl, "Martingale measures and hedging for discrete-time financial markets", *Math. Oper. Res.* **24**, 509-528, 1999.

[38] M. Schäl, "Portfolio optimization and martingale measures", *Mathematical Finance* **10**, 289-304, 2000.

[39] M. Schäl, "Price systems constructed by optimal dynamic portfolios", To be published in *Math. Methods Op. Res*, 2000.

[40] M. Schweizer, "Variance-optimal hedging in discrete time", *Math. Oper. Res.* **20**, 1-32, 1995.

[41] M. Schweizer, "Approximation pricing and the variance-optimal martingale measure", *Ann. Prob.* **24**, 206-236, 1996.

[42] S. E. Shreve and J. *Večeř*, "Options on a traded account: vacation calls, vacation puts and passport options", *Finance Stochast.* **4**, 255-274, 2000.